

The Research on Enterprise Micro-Blog Analysis Method Based on Knowledge Network

Wanjun Liu

School of Business Administration, South China University of Technology, Guangzhou, China

Email: 18664826331@163.com

How to cite this paper: Liu, W.J. (2017) The Research on Enterprise Micro-Blog Analysis Method Based on Knowledge Network. *Open Journal of Social Sciences*, 5, 127-138. <https://doi.org/10.4236/jss.2017.52013>

Received: January 3, 2017

Accepted: February 14, 2017

Published: February 17, 2017

Copyright © 2017 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Microblogging, as a typical representative of social networking applications, its advantages in interpersonal communication and information dissemination are very obvious. At present, micro-blog has become an important enterprise information dissemination channels and customer management tools. In this paper, based on the construction of enterprise microblogging knowledge network, combined with social network analysis, the author puts forward a set of micro-blog enterprises in-depth analysis of practical methods. It has a certain reference value for the further development of the value of enterprise microblogging data.

Keywords

Knowledge Network, Enterprise Micro-Blog, Social Network Analysis

1. Introduction

Microblogging has become an important platform for public opinion, media communication, corporate brand and product promotion because of the nature of sharing, interaction and openness. Enterprise microblogging is a kind of accounts that enterprise open in the microblogging platform [1] [2]. As an independent communication media on social media, because of its important commercial value, enterprise micro-blog has become an important network marketing means. So it is of great practical significance to study the micro-blog. In order to effectively mine the information in enterprise microblogging and give full play to the commercial value of enterprise microblogging, this paper presents a set of enterprise micro-blog analysis method based on knowledge network. In the second chapter, we briefly review the current research situation, construct the enterprise micro-blogging knowledge network model in chapter three, then propose the enterprise micro-blog analysis method based on knowledge network

in chapter four, and finally, in chapter 5, the feasibility of the model is verified by taking Club of Huawei as an example.

2. An Overview of Related Studies

Enterprise microblogging, an enterprise account in the microblogging platform, as an independent medium on social media, has very important commercial value. B. J. Jansen and M. Zhang, who analyzed 150,000 microblogging information in Twitter, found that 3.8% of the microblogging messages conveyed the brand's emotional tendencies, and that the micro-blog site to become the platform of enterprise marketing, customer relationship maintenance and word of mouth [3]. However, the length of microblogging text is short, the amount of words is small, and the use of words is not standardized, which brings great difficulty for effective analysis of enterprises microblogging.

The concept of knowledge networks was first proposed by the Swedish industry, the relevant research began in the 20th century, the mid-90s. Beckmann described the knowledge network as institutions and activities for production and dissemination of scientific knowledge in the academic point of view [4]. Andreas believes that the knowledge network is a social network among the participants. Participants at all levels realize the production and transfer of knowledge through this knowledge network, thus creating value [5]. These characteristics of knowledge network make it fit well with the originality, communication and interaction of micro-blog. Building enterprise microblogging knowledge network can dig out the inherent value of enterprise microblogging better.

3. The Construction of Enterprise Micro-Blog Knowledge Network

3.1. Microblogging Content Acquisition

At present, Sinamicroblogging API (Application Programming Interface) content is not fully open. There are many restrictions in the number of single-query results, access to data resources and call frequency. So simply using the API provided by Sinamicroblogging to obtain comprehensive microblogging business data is more difficult. The data acquisition efficiency is relatively low too. Therefore, this paper choose to use the Web crawler technology to obtain business microblogging data.

Web crawler is a client program, used to obtain information on the web page. How Web Crawlers Work:

- 1) Establishes a connection with the server, sends an Http request to the server, and requests a web page;
- 2) The server receives the request, make a response, and return to the web crawler page html source code;
- 3) Web crawler analysis html page, access to the URL contained in html, joined the URL queue;
- 4) If the URL is still not crawled in the URL queue, return to step 1 to continue.

3.2. Enterprise Micro-Blog Knowledge Network Node Access

Using the Web crawler to get a collection of microblogging content as follows:

$$M = \{(m_i) | i = 1, \dots, a\} \quad (1)$$

where i is the microblogging number, a is the total number of microblogs to obtain, M is the collection of all microblogging content.

In order to obtain the key words of each micro-blog text, we must first segment the micro-blog text, and then get the enterprise micro-blog node set K , as follows:

$$K = \{(k_j) | j = 1, \dots, n\} \quad (2)$$

The word frequency of k_j is expressed by $q(k_j)$, and the high frequency word set K_1 is established as follows:

$$K_1 = \{(k_j) | q(k_j) > q_1\} \quad (3)$$

q_1 is the threshold value, used to distinguish between high frequency and low frequency keywords.

3.3. Network Weights

In the enterprise microblogging, the keywords with high frequency often represent the focus of microblogging communication. And obtains frequency weight set Q of the high frequency keywords according to the word frequency of the acquired keywords:

$$Q(K_1) = \{q(k_j) | k_j \in K, q(k_j) > q_1, j = 1, \dots, n\} \quad (4)$$

In the enterprise micro-blog social network, the two keywords can be linked through microblogging content. The more the number of times two keywords appear in the same microblogging content, the closer the relationship between the two topics is [6]. According to whether the two high-frequency nodes in the same micro-Bo, we can build enterprise micro-blog high frequency node co-occurrence relationship set. The $e_{ij} = 1$ indicates that the high-frequency node k_i and the high-frequency node k_j co-occur in the same enterprise micro-blog, and $e_{ij} = 0$ indicates that there is no co-occurrence relationship between the nodes k_i and k_j . Assuming a total of n high frequency keywords, then the high frequency node co-occurrence relationship set can be expressed as:

$$E = \{(e_{ij}) | e_{ij} = 1, q(k_i) > q_1, q(k_j) > q_1; i, j = 1, 2, \dots, n\} \quad (5)$$

Based on the number of co-occurrence relations between high-frequency nodes, the co-occurrence relationship weight set of the enterprise micro-blog high-frequency nodes can be constructed as follows:

$$Q(E) = \{q(e_{ij}) | i, j = 1, \dots, n\} \quad (6)$$

The greater the value of $q(e_{ij})$, the greater the link between the two key-

words. Through the statistics the number of times that all the keywords appeared in the same micro-blog, we could be able to build enterprise micro-blog high frequency keywords co-occurrence matrix.

3.4. The Construction of Enterprise Micro-Blog Knowledge Network

According to the enterprise micro-blog high-frequency keyword set obtained from (3), high frequency key word frequency set obtained from (4) and high frequency node co-occurrence times obtained from (6), we can get enterprise micro-blog Weighted Knowledge Network Model as follows:

$$EMKN = \{K_1, Q(K_1), Q(E)\} \quad (7)$$

From the formula (7), we can know that the nodes have two kinds of weights: $Q(K)$ and $Q(E)$, where $Q(K)$ denotes the number of times the node appears, $Q(E)$ represents the number of co-occurrence between nodes. We can weigh the importance of the node from these two respects.

4. Analysis Method of Enterprise Micro-Blog Based on Knowledge Network

4.1. Identification of Important Nodes Based on Word Frequency

Using the EMKN model, the nodes with larger weights can be identified and analyzed. The high-frequency keyword set is composed of keywords whose word frequency is greater than a certain high-frequency threshold. This collection represents the point at which the enterprise wants to focus on disseminating knowledge to users in all published microblogs. The threshold of high-frequency keywords can be determined by the enterprise microblogging according to the actual situation.

4.2. Centre-Point Identification Based on Co-Occurrence Matrix

In the social network, the centrality can be used to measure the node's position in the overall network. The nodes with high centrality are at the core position in the whole network. Other nodes are either directly related to them, or connected to other nodes through them [6]. We refer to these nodes as "central points" and measure their importance with two points: degree centralization (C_D) and betweenness centralization (C_B) [7].

In the network, if the degree of a node is larger, it means that the higher the centrality of this node, the greater the importance of the node in the network.

Without considering the weight of the edge, the degree centralization of the node k_i is:

$$C_D(k_i) = \left[\sum_{j=1}^n e_{ij} \right] / (n-1) \quad (8)$$

Considering the weight of edge, the degree centralization of node k_i is:

$$WC_D(k_i) = \left[\sum_{j=1}^n q(e_{ij}) \right] / (n-1) \quad (9)$$

The significance of betweenness centralization can be expressed as follows: if two non-adjacent nodes s and t want to interact with each other and node i is on their path, the node i may control the interaction between them [8]. Therefore, in the constructed knowledge network, if the node i is located on the path that many nodes are connected with each other, the node i is in the most important position in the network.

Similarly, we can calculate the betweenness centralization $C_B(k_i)$ of K_i :

$$C_B(k_i) = \frac{2 \sum_{s < t} g_{st}(k_i) / g_{st}}{(n-1)(n-2)} \quad (10)$$

In the above formula, g_{st} denotes the number of shortcuts for nodes k_s to k_t , and k_i indicates the number of shortcuts to pass through k_i .

Based on the above centrality metric, we can find that the node with high centrality is very important for discovering the microblog topic information.

4.3. Cluster Identification Based on Cohesive Subgroups

Using the EMKN model, we can cluster the enterprise micro-blog by the method of cohesive subgroup analysis. At the same time, we can use the network midpoint and edge weights to carry on the further analysis. This way is not only able to analyze the weight of various groups of enterprise micro-blog and the relationship between groups, but also the composition of various groups. Importantly, the visualization of the analysis results helps to identify the key communication points of the enterprise micro-blog effectively. In addition, the knowledge points of any subgroup can be represented by a knowledge subnet. With this knowledge subnet, we can deeply analyze the internal structure, hot spots and associated patterns of sub-groups.

5. Examples and Application of Models

In this study, we selected the Club of Huawei, official microblogging of Huawei, as the research object in Sina micro-blogging platform. In recent years, Huawei's brand awareness and reputation have been greatly improved. The National Federation of Industry and Commerce released "2016 top 500 Chinese private enterprises" list, Huawei become the top 500 list with 395.09 billion Yuan in annual revenue. Club of Huawei is an interaction platform for Huawei's fans. It answers questions of fans, presents the latest product and service information, provides rich online content and offline interaction at the first time. Therefore, this paper selects this microblogging account for data collection.

5.1. Microblogging Access and Processing

First of all, the preparation of reptiles collects all the microblogging of Club of Huawei in January 1, 2015 to December 31, 2015, a total of 2974. Then segment the microblog text: we use the NLP/ICTCLAS 2014 Chinese word segmentation system to preprocess the crawled microblogging text, and get the keywords by all the microblogs. The third step is to obtain high-frequency keywords. In

this paper, we choose 104 words with frequency more than 50 as high frequency keywords. **Table 1** shows some of the high frequency words and their frequency.

5.2. Construction of Knowledge Network

According to the formula (4), we can get the word frequency of the high frequency keywords. Construct the matrix between microblogging content and keywords, form 2974×104 word matrix. In the enterprise microblogging text, there must be some kind of association between co-occurrence keywords, the degree of association can be expressed with the frequency of co-occurrence. According to formula (6), the co-occurrence matrix of 104 high-frequency keywords is obtained, which is 104×104 co-word matrix.

According to the co-occurrence relationship between high frequency keywords, the knowledge network model of Club of Huawei was constructed with Ucinet software, as shown in **Figure 1**.

Table 1. High frequency keywords and word frequency table (part).

Keyword	Freq.	Keyword	Freq.	Keyword	Freq.	Keyword	Freq.
Pollen	2455	Live micro-blog	339	Honor 71	238	Pollen annual meeting	194
Huawei	1137	Micro-blog	313	Entertainment	227	China	180
Mobile phone	738	Winning yesterday	267	Function	210	Global	177
Honor	681	Gift	264	Youth	199	Intelligence	165
Minister	509	Conference	256	World	195	Product	162
Topic	478	Huawei P8	238	Excellent Purchase Code	195	Headset	158

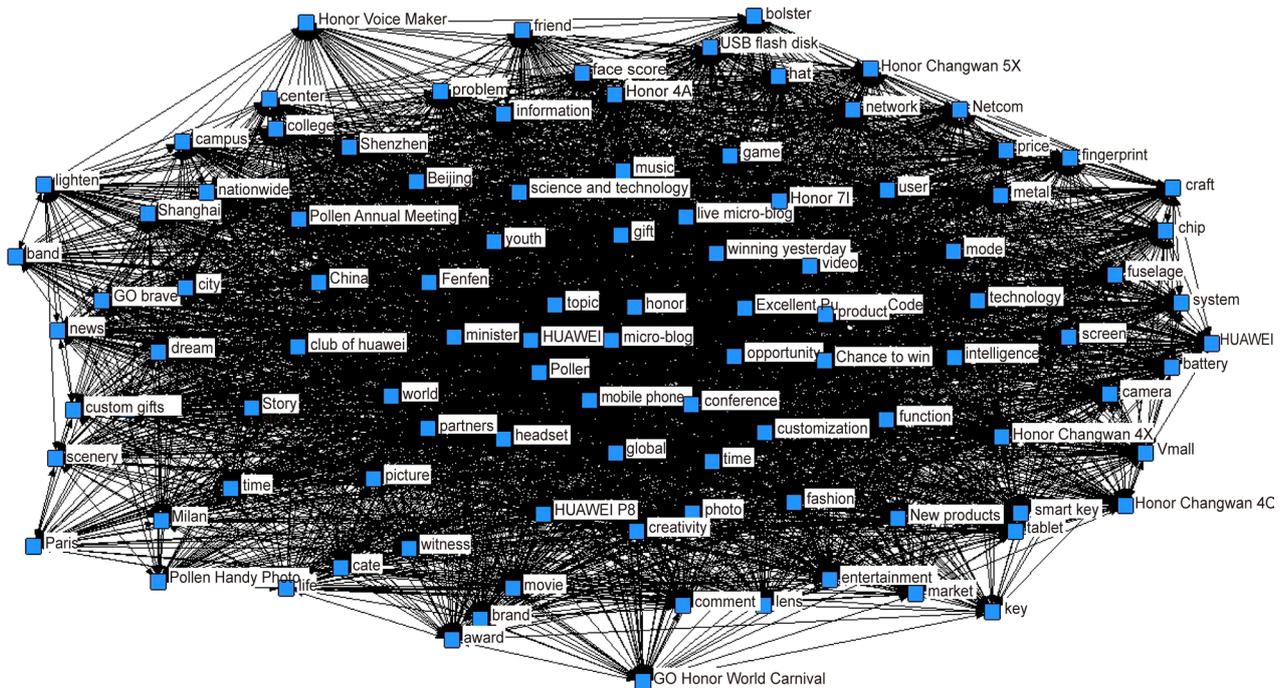


Figure 1. High frequency keywords co-occurrence network.

5.3. Identification and Analysis of Important Nodes

Enterprise microblogging, as a main positions of enterprise marketing and promotion and an important window of connecting with the user directly, its content often represents the main information enterprise wants to pass to its users. Therefore, in general, the high-frequency words appeared in the enterprise microblogging, often on behalf of the focus of the enterprise microblogging content.

We can see from **Table 2** that the words “Pollen”, “Huawei”, “mobile phone” and “honor” are the most frequently mentioned words in the microblog of Club of Huawei. From the above list of high-frequency keywords, we could found that Club of Huawei focused on Huawei’s mobile phone users, known as “Pollen”. In 2015, Huawei launched a number of different models of new mobile phones, including the honor series, mate series, P series, Maimang series and so on. But according to all the high frequency vocabulary released by Club of Huawei in 2015, we can found: In 2015, the Club of Huawei focused on promoting the honor series and “Huawei P8”. Although the Mate series is the focus of the launch in 2015, a high-end model, but its micro-blog marketing share is not significant.

5.4. Identification and Analysis of Centre Point

In order to further study the importance of the key nodes in the Club of Huawei microblogging, refers to the centrality of the social network, this paper uses the degree centralization and the betweenness centralization of node to identify the central point in the microblogging. The analysis results are shown below:

As can be seen from **Figure 2**, degree centralization of “pollen”, “honor”, “Huawei”, “topic”, “mobile phone” is relatively high, indicating that these keywords and other microblogging keywords directly related to more. However, comparing with the high frequency vocabularies, although the frequency of these keywords, such as “Honor 7i” and “Honor Changwan 5X”, is relatively high, but their degree centralization does not reach the equivalent height. This means that although the frequency of these keywords is relatively large, but their relevance with keywords appeared in other microblogging is not high. While both the frequency and the degree centralization of “Huawei P8” are relatively high, we can easily find that the model is the Club of Huawei focus in 2015 microblogging.

Table 2. High frequency keywords co-occurrence matrix (part).

	Pollen	Huawei	Mobile phone	Honor	Minister	Topic	Live micro-blog
Pollen	1724	323	224	475	251	221	134
Huawei	323	652	179	155	78	162	150
Mobile phone	224	179	452	180	50	71	42
Honor	475	155	180	1002	183	180	138
Minister	251	78	50	183	444	57	39
Topic	221	162	71	180	57	475	312
Live micro-blog	134	150	42	138	39	312	339

The betweenness centrality directly reveals the importance of the node location in the whole knowledge network. As we can see from **Figure 3**, the betweenness centralization of “pollen”, “Huawei”, “mobile phone” and “glory” are relatively high. These nodes are located at the core of the network and play an important mediating role.

It is noteworthy that although the frequency and degree centralization of “Huawei P8” are relatively high, its betweenness centralization is only 17.241,

		1 Degree	2 NrmDegree	3 Share
1	Pollen	6336.000	13.077	0.095
4	honor	4308.000	8.892	0.064
2	HUAWEI	3091.000	6.380	0.046
6	topic	2672.000	5.515	0.040
3	monile phone	2536.000	5.234	0.038
5	minister	1991.000	4.109	0.030
7	live micro-blog	1884.000	3.889	0.028
9	winning yesterday	1849.000	3.816	0.028
10	gift	1810.000	3.736	0.027
8	micro-blog	1555.000	3.209	0.023
11	conference	1186.000	2.448	0.018
18	Excellent Purchase Code	1164.000	2.402	0.017
24	headset	988.000	2.039	0.015
12	HUAWEI P8	912.000	1.882	0.014
15	function	888.000	1.833	0.013
21	global	808.000	1.668	0.012
23	product	797.000	1.645	0.012
32	Chance to win	786.000	1.622	0.012
44	customization	777.000	1.604	0.012
17	world	773.000	1.595	0.012
19	Pollen Annual Meeting	728.000	1.503	0.011
16	youth	715.000	1.476	0.011
30	opportunity	695.000	1.434	0.010
22	intelligence	684.000	1.412	0.010

Figure 2. Node degree centrality (partial).

		1 Betweenness	1 nBetweenness
1	Pollen	35.638	0.692
2	HUAWEI	35.638	0.692
3	monile phone	33.702	0.654
4	honor	33.148	0.644
6	topic	32.796	0.637
5	minister	32.771	0.636
8	micro-blog	32.277	0.627
18	Excellent Purchase Code	27.099	0.526
9	winning yesterday	26.877	0.522
30	opportunity	26.869	0.522
11	conference	26.586	0.516
10	gift	26.392	0.512
16	youth	26.259	0.510
32	Chance to win	26.077	0.506
7	live micro-blog	25.971	0.504
17	world	25.731	0.500
21	global	25.152	0.488
27	Fenfen	24.500	0.476
42	video	24.263	0.471
39	partners	23.832	0.463
15	function	23.612	0.458
23	product	23.567	0.458
13	Honor 71	21.992	0.427
40	Photo	21.195	0.411
44	customization	20.824	0.404
52	Beijing	20.692	0.402

Figure 3. Node betweenness centrality (partial).

which means that it is not very important in the whole network. This is because, as an independent high-end model in addition to the honor series, the betweenness role of “Huawei P8” is not obvious.

5.5. Identification and Analysis of Cluster

By analyzing the cohesive subgroup of the enterprise microblogging knowledge network, it is possible to divide the microblog topic. Analysis results into the following:

As shown in **Figure 4**, 103 high frequency keywords are divided into 8 subgroups by the cohesive subgroup analysis. The composition of each subgroup is as follows:

It can be seen from **Table 3** that Club of Huawei has different marketing methods for different models. Specific analyses of various types of groups are as follows:

Group 1 mainly includes some topics and activities to increase the stickiness of microblogging fans, such as the activities of “Pollen Handy Photos”, “food”, and “entertainment”. These microblogging related to the daily life of fans can effectively increase the fan’s activity.

Group 2 mainly includes some marketing activities of “Honor Changwan 5X”. According to the composition of the key words group 2, we can find that, for this product, microblogging marketing activities are excellent purchase code and lottery mainly.

Group 3 mainly includes marketing activities of “Huawei P8”. This type of group contains the most keywords, we can see, for this high-end model, the marketing activities are the most diverse.

Group 3 mainly includes marketing activities of “Huawei P8”. This type of group contains the most keywords, we can see, for this high-end models, the marketing activities are the most diverse.

Group 4 mainly includes marketing activities of new product of honor series in 2015. From the composition of group 4, we could found that the Club of Huawei launched a number of marketing activities for college students. Therefore, we can easily draw the conclusion that the products of honor series mainly targeted at young users, especially college students. This is also very consistent with the low-end machine positioning of honor series products.

Group 5 mainly includes “tablet”, “Honor Changwan 4C”, “Huawei Changxiang” and so on. By the group 5, we can know that these products are not Club of Huawei microblogging marketing focus, their marketing activities are also very single, and “Price”, “brand” are their main selling point.

Group 6 is the product information most relevant to the actual use of the user, including “key”, “lens”, “video”, “information”, and “function”.

Group 7 is the topic of product design, including the “battery”, “craft”, “metal”, and “fuselage”.

Group 8 mainly includes keywords related to product features. From these key words, we can easily find that Huawei mobile phones in 2015 mainly has made a

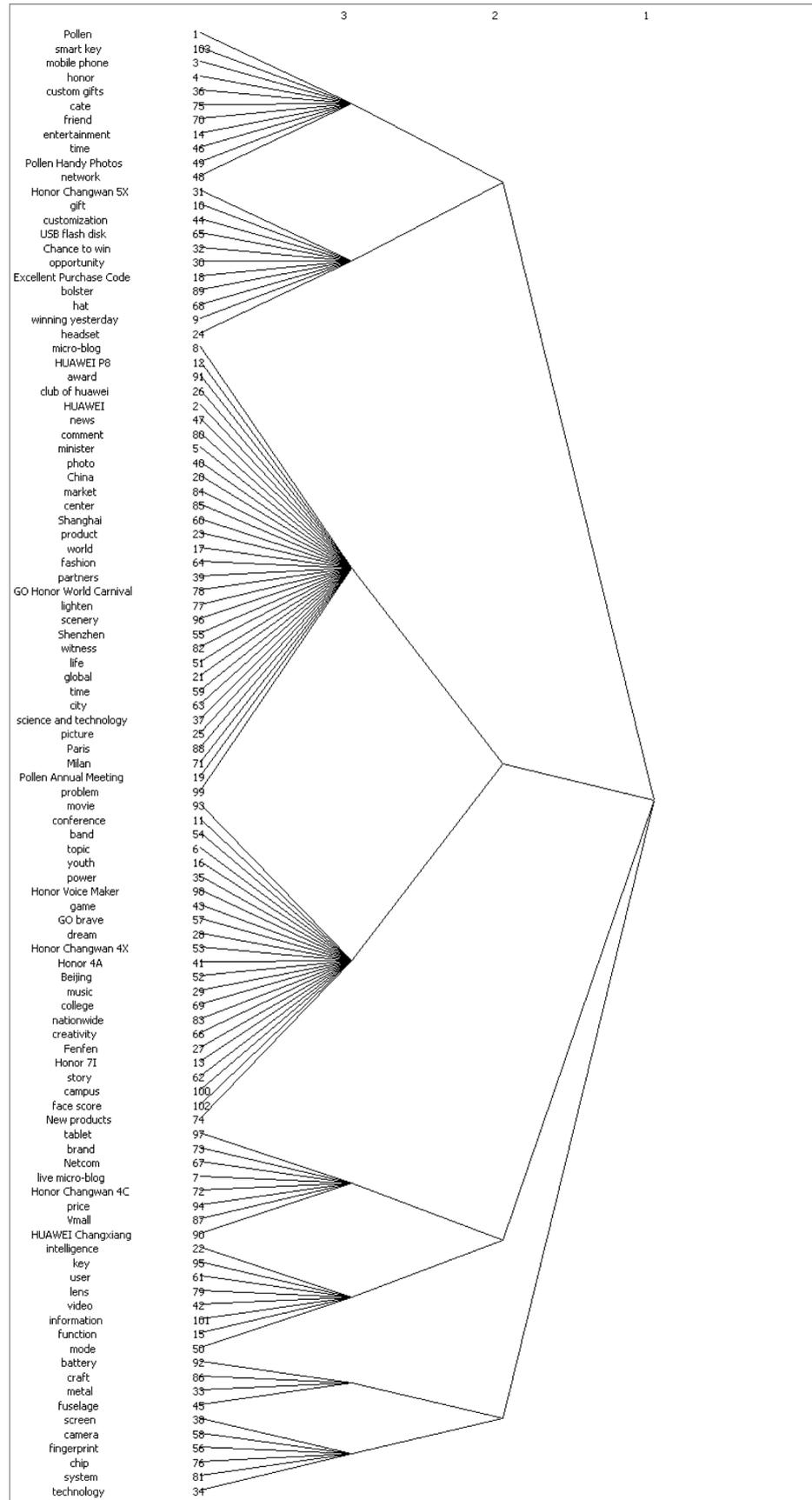


Figure 4. The chart of cohesive subgroups of high frequency key words.

Table 3. High-frequency keyword classification table.

Cluster	Keywords
1	Pollen (Huawei's fans), smart key, mobile phone, honor, custom gifts, food, friend, entertainment, time, Pollen Handy Photos, network
2	Honor Changwan 5X, gift, custom, USB flash disk, Chance to win, opportunity, Excellent Purchase Code, bolster, hat, winning yesterday, headset
3	micro-blog, Huawei P8, award, Club of Huawei, Huawei, news, comment, minister, photo, China, market, center, Shanghai, product, world, fashion, partners, GO Honor World Carnival, lighten, scenery, Shenzhen, witness, life, global, time, city, science and technology, picture, Paris, Milan, Pollen Annual Meeting, problem
4	movie, conference, band, topic, youth, power, Honor Voice Maker, game, GO brave, dream, Honor Changwan 4X, Honor 4A, Beijing, music, college, nationwide, creativity, Fenfen (The nickname of pollen), Honor 7I, story, campus, face score, new products
5	tablet, brand, Netcom, live micro-blog, Honor Changwan 4C, price, Vmall, Huawei Changxiang
6	intelligence, key, user, lens, video, information, function, mode
7	battery, craft, metal, fuselage
8	screen, camera, fingerprint, chip, system, technology

breakthrough in the “screen”, “camera”, “fingerprint”, “chip”, “system”, and “technology”.

6. Conclusion

This paper constructs a micro-blogging knowledge network with micro-blog text keywords as network nodes, keyword frequency and the co-occurrence relation between them as weights and edges and effectively found the important nodes, the central nodes and the categories of the microblogging transmission through the central analysis of the micro-blogging knowledge network and the cohesive subgroup analysis. Then, with Huawei's official micro-blog “Club of Huawei” as the research object, this paper constructs the knowledge network of Club of Huawei, and effectively analyzes its key products, marketing selling points and main marketing activities, and further analyzes the major improvement in the field of mobile phone products in 2015 in Huawei. Through the example, we found that the knowledge network model can carry on the thorough and comprehensive analysis to the enterprise's micro-blog, which is of great significance to the micro-blog marketing and the enterprise's competitive intelligence.

References

- [1] Riemer, K., Richter, A. and Seltsikas, P. (2010) Enterprise Microblogging: Procrastination or Productive Use? *AMCIS*, 506.
- [2] Kaplan, A.M. and Haenlein, M. (2011) The Early Bird Catches the News: Nine Things You Should Know about Micro-Blogging. *Business Horizons*, **54**, 105-113. <https://doi.org/10.1016/j.bushor.2010.09.004>
- [3] Jansen, B.J., Zhang, M., Sobel, K., *et al.* (2009) Twitter Power: Tweets as Electronic Word of Mouth. *Journal of the American Society for Information Science & Tech-*

- nology*, **60**, 2169-2188. <https://doi.org/10.1002/asi.21149>
- [4] Beckmann, M.J. (1995) Economic Models of Knowledge Networks. In: Batten, D. and Casti, J., Eds., *Networks in Actions*, Heidelberg, Berlin, Springer-Verlag, New York, Tokyo, 159-174. https://doi.org/10.1007/978-3-642-57843-4_7
- [5] Seufert, A., Krogh, G.V. and Bach, A. (1999) Towards Knowledge Networking. *Journal of Knowledge Management*, **3**, 180-190. <https://doi.org/10.1108/13673279910288608>
- [6] Liu, J. (2004) *Social Network Analysis*. Social Sciences Academic Press (CHINA), Beijing.
- [7] Liao, X., Li, Z.H. and Xi, Y.J. (2016) Modeling and Analyzing Methods of User-Innovation Knowledge in Enterprise Communities Based on Weighted Knowledge Network. *System Engineering—Theory & Practice*, **36**, 94-105.
- [8] Liu, J. (2009) *Lectures on Whole Network Approach*. Shanghai People's Publishing House (SPPH), Shanghai.



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact jss@scirp.org